



Majid, A., Jordan, F. M., & Dunn, M. (2015). Semantic systems in closely related languages. *Language Sciences*, 49, 1-18.
<https://doi.org/10.1016/j.langsci.2014.11.002>

Peer reviewed version

Link to published version (if available):
[10.1016/j.langsci.2014.11.002](https://doi.org/10.1016/j.langsci.2014.11.002)

[Link to publication record in Explore Bristol Research](#)
PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Semantic systems in closely related languages

Asifa Majid

Radboud University, Nijmegen, The Netherlands

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Fiona Jordan

University of Bristol, United Kingdom

Michael Dunn

Uppsala University, Sweden

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Abstract

In each semantic domain studied to date, there is considerable variation in how meanings are expressed across languages. But are some semantic domains more likely to show variation than others? Is the domain of space more or less variable in its expression than other semantic domains, such as containers, body parts, or colours? According to many linguists, the meanings expressed in grammaticised expressions, such as (spatial) adpositions, are more likely to be similar across languages than meanings expressed in open class lexical items. On the other hand, some psychologists predict there ought to be more variation across languages in the meanings of adpositions, than in the meanings of nouns. This is because relational categories, such as those expressed as adpositions, are said to be constructed by language; whereas object categories expressed as nouns (e.g., containers) are predicted to be “given by the world”. We tested these hypotheses by comparing the semantic systems of closely related languages. Previous cross-linguistic studies emphasise the importance of studying diverse languages, but we argue that a focus on closely related languages is advantageous because domains can be compared in a culturally- and historically-informed manner. Thus we collected data from 12 Germanic languages. Naming data were collected from at least 20 speakers of each language for containers, body-parts, colours, and spatial relations. We found the semantic domains of colour and body-parts were the most similar across languages. Containers showed some variation, but spatial relations expressed in adpositions showed the most variation. The results are inconsistent with the view expressed by most linguists. Instead, we find meanings expressed in grammaticised meanings are more variable than meanings in open class lexical items.

Keywords: semantics, evolution, colour, body parts, containers, spatial relations, Germanic languages

Introduction

Why study word meaning?

When contemplating the uniquely human aspects of language, researchers typically focus on grammar, an important feature of language without doubt. Far less effort, however, is expended on theorizing about meaning, especially those aspects of meanings encoded in the lexicon. Other species have rudimentary (although impressive) abilities to understand words. Rico the border collie is reported to know the names for 200 items (Kaminski, Call, & Fischer, 2004); while Alex the Grey parrot knew 50 words for objects, 7 for colours, 5 for shapes, and 8 for quantities (Pepperberg, 2012). But contrast this with the average English-speaking child who at 4 years already has in her repertoire 5,000 words. By 8 she knows 10,000 words, and as an adult she has a vocabulary of around 20,000-35,000 words. In fact, word learning continues into middle age (where it seems to plateau), with speakers adding almost 1 new word to their vocabulary every day (R.L.G., 2013).

Despite this incredible feat, particular to our species, we know astonishingly little about this aspect of language. From Bloomfield to Chomsky, the same sentiment echoes: “The most elaborate dictionaries provide no more than bare hints about the meanings of words” (Chomsky, 1993, p. 23); or “The statement of meanings is therefore the weak point in language-study” (Bloomfield, 1933, p. 140). Summarising the progress in the field Weinreich (1980, p. 268) concludes: “... the mushrooming literature of semantics has been strangely noncumulative, mainly because in contrast to phonology or grammar, the gathering of facts and theoretical speculation have not, in semantics, gone hand in hand”.

This paper introduces a new endeavour: The “Evolution of Semantic Systems” project. Here theoretical perspectives from psychology, linguistics, and anthropology, are joined hand-in-hand with a large-scale empirical study of lexical semantics across a variety of domains. We aim to address some perennial questions, such as: How variable are semantic categories across languages? What factors predict similarities and differences in meaning? Are some domains more variable than others? What variation is there in the patterns and processes of historical change for different semantic categories? Within the Evolution of Semantic Systems project, or EoSS for short, we investigate how meanings vary over space, and change over time. We focus on different kinds of categories: containers (kinds of objects), colour (attributes of objects), body parts (parts of objects), and spatial relations (how objects are related to one another). More on each of these domains follows later.

Previous studies of semantics often focus on the meaning of individual words. They also primarily rely on meanings that can be extracted from texts. These investigations shed light on general processes such as how words get to have broader or narrower meanings (generalisation/specialisation), how a part begins to stand for the whole (metonymy), how words attain positive or negative connotations (melioration/pejoration), and so forth. Our project employs a different perspective. We investigate how words in lexical fields are used to categorise entities of various sorts. By using an objective referential grid of comparison across languages we can quantify similarity, difference, and change across languages. We focus specifically on extensional semantics, i.e., how similar, or different the referential ranges (“boundaries”) of words are across languages. For example, does *gelb* ‘yellow’ in German cover the same hues as *geel* in Dutch? Does the meaning of *arm* across languages include or exclude the ‘hand’? This does not mean we are equate meaning with extension; or that we are ignore intension. We take it as given that an intensional analysis is one of the ultimate goals of a semantic analysis.

Nevertheless we seek empirical data on the variation and similarity across domains in a comparable framework, and believe this is a crucial part of the data an intensional analysis will ultimately be accountable to.

Methodologically, we take a phylogenetic approach, integrating into our analyses a class of computational techniques developed in evolutionary biology to deal with hierarchically related entities such as species or languages (Felsenstein, 2004; Nunn, 2011). The phylogenetic approach to variation further distinguishes this project from previous cross-linguistic studies of meaning, which have advocated using a sample of diverse languages from around the world as a way of minimising similarities due to shared language ancestry (e.g., Levinson, Meira, & The Language and Cognition Group, 2003; Majid, Bowerman, van Staden, & Boster, 2007; Majid, Enfield, & van Staden, 2006). However, little is known about what constitutes “diverse” for semantics. Moreover, existing studies typically over-represent easily accessible, national languages with a long literary history, and take a brute-force approach to control for language relatedness (i.e., they simply exclude related languages from their samples). The EoSS project addresses these issues by studying languages in the single well-resolved Indo-European family, where rigorous quantitative phylogenetic analyses of language relationships are available (Bouckaert et al., 2012). By tracking form and meaning in a single language family (or sub-family), we can begin to understand the evolutionary dynamics of semantic change. While this broader phylogenetic work is in progress, this volume presents intermediate results from the Germanic sub-family, i.e., a precursor to quantitative phylogenetic analyses. In this paper, we present the theoretical background to the EoSS project, and some initial outcomes in order to exemplify the general approach.

The rise of interest in meaning

“Semantics is a perfectly respectable linguistic activity.” (Bazell, 1953, p. 89)

In the last decade, there has been a burst of empirical studies investigating semantics cross-linguistically (see Majid, 2014, and Malt & Majid, 2013, for overviews). The model for many of these studies is the pioneering work of Berlin and Kay (1969). In their study of colour, Berlin and Kay introduced a methodology for systematically collecting data from different languages which was further refined in the World Colour Survey (Kay, Berlin, Maffi, Merrifield, & Cook, 2009) – the largest ever empirical study of semantics, featuring 110 languages spoken primarily by small-scale, preliterate, non-industrialised communities.

The starting point for this survey was a non-linguistic stimulus set that served as a prompt for linguistic elicitation. Speakers of different languages were presented with colour chips one at a time, and the names given to these colours indexed the referential range of the terms. By examining the extension of colour terms across languages, Kay and colleagues were able to measure the similarity in meaning across languages (e.g., Kay & Regier, 2003; Regier, Kay, & Khetarpal, 2007). The upshot of these studies was that the semantics of colour terms are not random across languages (Kay & Regier, 2003), but reflect cognitive and communicative principles (e.g., Regier, Kay, & Khetarpal, 2007; Steels & Belpaeme, 2005). The World Colour Survey also revealed there are important differences between languages: some languages have three terms, others eleven (or even fifteen; cf. Roberson, Pak, & Hanley, 2008). Nevertheless, across diverse languages boundaries and foci for colours are orderly. Building on this database, other researchers have attempted to account for the apparent variation, appealing either to the physical environment (e.g., Lindsey & Brown, 2004) or cultural practices (e.g., Levinson, 2000).

It is important to keep in mind that Berlin and Kay’s work has been an inspiration not just for other large-scale quantitative studies, but also for particularistic studies of individual

languages (al-Jehani, 1990; Davidoff, Davies, & Roberson, 1999; de Vos, 2011; Hill, 2011; Levinson, 2000; Lindsey, Brown, Brainard, & Apicella, 2014; Turton, 1980). In juxtaposition to the abstraction required for large-scale cross-linguistic comparison, fine-grained case studies provide critical contextualization, and offer alternative modes of understanding the data. The two go hand-in-hand: We can appreciate the differences by understanding the similarities.

Surprising perhaps, against this background, is the fact that there is little comparative empirical data on colour for the Indo-European language family. Berlin and Kay's original work is often criticised for over-sampling Indo-European languages (e.g., Lucy, 1997; Wierzbicka, 2005). From a different standpoint they could be said to under-sample them, as Berlin and Kay only present primary colour data for a handful of Indo-European languages: (American) English, (Mexican) Spanish, Catalan, Bulgarian, and Urdu. It is intriguing to see, but less often remarked upon, that there are notable differences between these related languages. For example, Urdu only has eight basic colour terms, whereas the other languages appear to have eleven. Berlin and Kay also briefly discuss the possibility that Russian has twelve basic colour terms, depending on the exact status of Russian's two 'blues' *siniy* and *goluboy* (Berlin & Kay, 1969, p.35-36). Thus Berlin and Kay's work could be taken to show notable differences in colour semantics within Indo-European languages.

Since Berlin and Kay, there have been isolated studies of colour terminologies in specific languages, and sub-families such as Celtic (Lazar-Meyn, 1991) and Slavonic (Comrie & Corbett, 1993), but these studies are (understandably) limited in scope. Perhaps most remarkably there has been no published study of colour terms in the Germanic language family (until this one).

The World Colour Survey was proof of concept that comparable data can be collected and compared across languages. It was also a major source of inspiration for studies to come in other

semantic domains. Three of these are particularly relevant in the context of our project: the human body and its parts, spatial relations, and containers.

Following Berlin and Kay's *Basic Colour Terms*, two scholars (Andersen, 1978; Brown, 1976) attempted their own typology of body part terms, and were able to articulate some generalisations, for example: a label for "foot" implies a label for "hand" (Andersen, 1978, p. 352); and if both "hand" and "foot" are labelled, they have different (and simple) names (Brown, 1976, p. 405). These generalisations, although thought-provoking, are problematic. They assume that speakers of all languages share the same understanding of the terms glossed as "hand" and "foot". This assumption has been questioned more recently.

In a series of studies, Majid and colleagues investigated the cross-linguistic variability of referential meaning for body part terms, analogous to investigations of colour (e.g., Majid, Enfield, & van Staden, 2006; Majid & van Staden, 2015; Majid, 2010). They found languages differ in what parts get singled out for naming, and in the precise extensions those parts have. For example, Dutch and Japanese both have separate terms for hand and arm. Nevertheless, when asked to colour the native equivalent of "arm" on a line drawing of a body, half the Dutch speakers included the hand; but Japanese speakers were much less likely to include the hand. In contrast, when asked to colour the "hand", Dutch speakers always coloured from fingertips to wrist, whereas Japanese speakers extended the equivalent term to the elbow, or even the shoulder. This suggests that even such fundamental domains as the body can vary in their semantic encoding. The extent of this variation within a language-family is not known. We provide some critical data on this matter in the results section.

Next we turn to spatial relations, and the influential work of Bowerman and colleagues on topological spatial relations, i.e., notions such as containment, attachment, support, encirclement, adhesion, etc. (Bowerman, 1992; Bowerman & Pederson, 1992, 1993). Bowerman and Pederson

(1992) created a stimulus set, a series of line drawings that depicted a Figure and Ground object in different spatial configurations (i.e., support, containment, etc.). In response to a prompt question (e.g., *Where is the X?*), speakers of various languages were asked to describe the location of a Figure object in each picture (see also Levinson & Wilkins, 2006).

In a preliminary report, Bowerman and Choi (2001) presented a synthesis of the results from speakers of 38 different languages sampled from 25 different language families. They showed that the extension of terms over spatial situations followed an orderly principle. By tracing the extension of the terms used for different prototypical scenarios (e.g., support, i.e., “cup on table”; containment, i.e., “apple in bowl”), Bowerman and colleagues were able to show there was a unidimensional implicational hierarchy that could account for the data from all languages. For example, English speakers used the prepositions *on* for “cup on table” and “handle on door”, but *in* for “apple in bowl”. On the other hand, Berber speakers had a distinct term *x* for “cup on table”, but used *di* for both “handle on door” and “apple in bowl”. Meanwhile, Dutch speakers used *op* for “cup on table”, *aan* for “handle on door”, and *in* “apple in bowl”.

Although there was an ordered system to how spatial terms were used, Bowerman stressed the fact that not all spatial scenarios were treated as equivalent across languages (as the above English, Berber, and Dutch illustrate), and concluded: “what a language *counts* as (sufficiently like) support or (sufficiently like) containment is not given by the structure of reality or our perception of it, but is determined to a large extent by language-specific conventions for how to construe spatial scenes” (Bowerman & Choi, 2001, p. 487, emphasis their own). This was later echoed by Levinson, Meira, and The Language and Cognition Group (2003; cf., Regier, Khetarpal, & Majid, 2013). Whether these conventions are language-specific, or they hold more generally across language-families, is a question the EoSS project has begun to answer. We return to this in the results section.

Variation is evident in the final domain we consider here too, i.e., containers. In a series of studies Malt and colleagues have investigated the relationship between perceived similarity between containers and the names given to them (Ameel, Storms, Malt, & Sloman, 2005; Malt, Sloman, & Gennari, 2003; Malt, Sloman, Gennari, Shi, & Wang, 1999; Pavlenko & Malt, 2013). In a first study (Malt et al. 1999), English, Spanish, and Mandarin speaking participants were tested. People were shown colour photographs of bottle, jars, and other similar containers (e.g., milk carton, peanut butter jar, baby powder dispenser), and simply asked to name them. Not only did the speakers of the three languages differ in how many terms they used to name stimuli (5 for Mandarin, 7 for English, 15 for Spanish), but the extension of the terms also varied. Comparable variation was evident when Belgian Dutch and French speakers named dishes and containers (Ameel et al., 2005), and when English and Russian speakers named drinking containers (Pavlenko & Malt, 2013).

These results are particularly interesting because it has been proposed that the naming patterns for concrete objects are similar across languages (e.g., De Groot, 1993; Gentner & Boroditsky, 2001; Gentner, 1981) (we return to this below), but Malt and colleagues challenge this assumption. Thus, Malt et al. (1999, p. 230) concluded from their English-Spanish-Mandarin study: “the three languages show *substantially* different patterns of naming”, and Ameel et al. (2005, p. 75): “the analysis of similarities among naming distributions... revealed *substantial* differences between the naming patterns of French- and Dutch-speaking monolinguals” (emphasis added to both quotes). This is striking, indeed, because these studies are comparing closely related languages (English and Spanish; Dutch and French).

In general, this raises the question of what actually constitutes “substantial differences”. And, more importantly, do semantic systems across languages differ by the same magnitude across different domains?

Comparing across domains

As the previous discussion has illustrated, researchers studying different domains emphasise different aspects of the data: some focus on the similarities; some on the differences. As will also have become clear, all domains have points of *both* similarity and differences. But are some domains relatively *more* similar across languages in their semantic structure than others?

Within psychology, it has been argued that naming for objects is less cross-linguistically variable than naming for relational entities (Gentner, 1981; Gentner & Boroditsky, 2001). According to Gentner's framework some categories are shaped more by perceptual-cognitive principles and others more by language-internal principles. Thus concrete nouns – terms for objects (such as containers) – are said to be given to us by perception-cognition: the world presents them to us as individuated entities, and we simply have to learn our language's label for them. At the extreme other end, closed-class terms (e.g., conjunctions or determiners) are said to be determined entirely by language-internal factors, and so aspects of language history, or culture, might affect the structure of these entities. Close to these purely linguistically structured entities sit the prepositions used for describing spatial configurations. Although they are also subject to cognitive factors, according to Gentner, language-internal factors play a primary role in constituting their meaning.

This contrast between cognitive- and linguistic-dominance becomes important when predicting which domains should show more cross-linguistic variability. According to Gentner and Boroditsky (2001, p.217): “Relational systems vary more crosslinguistically than do object meanings”. This predicts that terms for containers should be more similar in meaning than terms for spatial relations. This prediction is also made by other scholars. For example, on reviewing work on the bilingual lexicon, De Groot (1993, p. 40) claims: “the translations of concrete words

share more of their representation than the translations of abstract words”. This is because: “concrete words refer to entities whose function is likely to be the same across languages” (De Groot, 1993, p. 41). This line of reasoning also suggests that containers should share more similarity in meaning than spatial relations.

What about other sorts of entities? Objects and relations are important ontological kinds, but where do properties or parts fit in? Should colour and body parts pattern more like containers or more like spatial relations? On the one hand, both colour and body parts are open class, and so both domains should show more similarity of meaning than spatial prepositions. Does this mean they should show the same amount of similarity as containers? Both domains could be hypothesised to be given to us by perception, and so one could predict colour and body parts ought to behave just like object categories. However, an important principle in the Gentner dominance hierarchy is not just whether entities are given by perception but whether they are individuated. Objects – whether naturally occurring or manufactured by humans – are individuated entities *par excellence*. In contrast, colour is a continuous perceptual field in which boundaries are not obvious. So, on this logic, colour terms should be less similar in meaning than container terms.

What about body parts? Are they close to containers? Perhaps these could be thought of as individuated entities. Some parts, at least, can be segmented at joints, or along other salient visual (de Vignemont, Tsakiris, & Haggard, 2005) or other sensory principles (e.g., de Vignemont, Majid, Jola, & Haggard, 2009). On the other hand, parts could be considered relational: after all, a part is defined against a whole. Taken together these arguments suggest body part terms should show more variation than container terms, but perhaps less variation than spatial relation terms.

To summarise, according to the Gentner framework, the domain of containers should show most similarity in meaning across languages, followed by colour and body parts (in no clear order), followed finally by spatial relations.

Not everybody would make this prediction, however. Many linguists hold the opposite view to that implied by the Gentner hierarchy. This follows from the canon that there is a restricted set of meanings encoded in closed class items (e.g., Haspelmath, 2003; Landau & Jackendoff, 1993; Talmy, 1983). So, Talmy, for example, notes that only certain domains, such as space, time, and causation, are coded in closed class terms, and further: “they are not free to express just anything within these conceptual domains, but are limited to particular aspects and combinations of aspects, ones that can be thought to constitute the “structure” of these domains” (Talmy, 1983, pp. 227–228). Along similar lines, Landau and Jackendoff (1993) in their seminal paper observe there are only 80 to 100 prepositions in English, while there are tens of thousands of nouns. From this they go on to argue: “there are so few prepositions because the class of spatial relations available to be expressed in language – the notions prepositions can mean – is extremely limited” (p.224).

In contrast to the meanings of closed class terms, many linguists believe meanings encoded in open class words are more malleable, and subject to fluctuating social and technological factors. It is not hard to find such examples. When Europeans invaded the Americas they took with them new technologies; some languages, such as Imbabura Quechua, borrowed the associated words for them, such as *arrusa* ‘rice’, *riluju* ‘clock’, and *simana* ‘week’ (from Spanish) (Haspelmath, 2009). Or consider the legacy on the English language of the British colonisation of India. It gave rise to innumerable new words, such as: *curry*, *toddy*, *veranda*, *cheroot*, *loot*, *compound*, *chintz*, *calico*, *gingham*, *shawl*, *bamboo*, *pagoda*, *typhoon*, *monsoon*, *mandarin*, *dingy*, *betel*, *mango*, *cheroot*, *pariah*, *teak*, etc. (Yule & Burnell, 1903). These sorts of

examples suggest that the meanings encoded in open class words are subject to varying forces, and so their meanings likewise diverge cross-linguistically. Perhaps this is most aptly summarised by Bloomfield's dictum: "The lexicon is really an appendix of the grammar, a list of basic irregularities" (Bloomfield, 1933, p. 274).

Based on this line of reasoning, we could make the opposite prediction to the Gentnerian one: languages should be more similar to one another for the meanings of spatial prepositions than for the meanings of container, colour, or body part words.

Testing cross-linguistic hypotheses

The usual approach to studying questions that call for cross-linguistic evidence is to take as diverse a sample of languages as possible. As indicated earlier, there is little consensus as to what counts as diverse. Linguistic typologists argue amongst each other over what constitutes an adequate sample size, and about the precise criteria for selecting a sample (e.g., Dahl, 2008; Dryer, 1989; Nichols, 1992; Perkins, 1989; Rijkhoff, Bakker, Hengeveld, & Kahrel, 1993). There is broad agreement that genealogical relationships must be taken into consideration, and for good reason: related languages may just be replicates of the same historical process. For example, if ten languages share a feature, we are less likely to conclude that feature is a universal property of all languages if those languages are related than if they come from diverse stock. But scholars disagree as to whether time-depth of language-relatedness should also be taken into account in their sampling method.

Controlling for language relatedness through sampling has the consequence that many languages have to be excluded from study because they come from large families. Thus, vast numbers of Austronesian languages (1,222 languages) or Niger-Congo languages (1,526

languages)¹ would simply be discarded under most sampling methods. However, this is not the only option. We can take this design “flaw” and make it into a “feature”. That is, we can exploit the fact that languages have shared histories, and utilise a controlled comparison approach, as has become practice within evolutionary anthropology (Fortunato & Jordan, 2010; Mace, Holden, & Shennan, 2005; Mace & Pagel, 1994).

There are two benefits in taking a controlled comparison approach. First, within a language family, speakers are likely to have a shared background, so we can be more certain that the semantic domains we study are functionally comparable in a culturally- and historically-informed way. Second, the formal expression of semantic elements is more constrained than in a diverse sample (e.g., all Indo-European languages arguably have an adposition class). Finally, establishing how much variation there is within one language family can be a step towards quantifying the kind of sample that would be needed for a maximally diverse sample. So this approach has utility for core linguistic questions (e.g., Dunn, Greenhill, Levinson, & Gray, 2011). If semantic categories show much diversity, then we can relax our criteria for sampling within language families, but if semantics shows much stability within language family then this has consequences for the potential size and scope of languages needed for a broader survey.

Here, we focus on the Indo-European language family, and more specifically the Germanic sub-branch. Since William Jones’ original proposal of the Indo-European language family (Jones, 1798), there has been intensive interdisciplinary study of these languages. This makes this language family particularly interesting for a study of semantics, as there is a rich body of historical and textual data to draw upon. Moreover, there are known points of contact and

¹ Numbers come from Ethnologue (Lewis, Simons, & Fenning, 2014).

influence from other language families and isolates which makes future investigations of the impact of these factors on lexical distinctions eminently tractable.

To return to the overarching question – *Are some semantic domains more similar than others across languages?* – we focus on a comparison of 12 Germanic languages. For the reasons just given, we believe a comparison of these languages is ideal. The communities share many cultural commonalities deriving from a shared history of Western European social and political interaction, and it is precisely this feature that enables us to ask whether, and where, differences exist. Do semantic systems diverge more for containers or spatial relations; colours or body parts? Or are all semantic domains under the same pressures and influences?

Methods

Languages and Participants

We collected data from 12 Germanic languages with the cooperation of a distinguished team of language experts. See Table 1. The Germanic language family had three main branches, one of which is now extinct (Eastern Germanic) and so does not feature in our sample. The Northern Branch consists of the Scandinavian languages (the first 5 languages in Table 1); the Western Branch comprises the remaining languages, including English, German, and Dutch.

Language	Country (of data collection)	Number of participants (female)	Mean age (and range)	Researchers
Icelandic	Iceland	21 (10)	29 (22-57)	Matthew Whelpton Þórhalla Guðmundsdóttir Beck
Faroese	Faroe Islands	20 (11)	18.7 (18-20)	Zakaris Hansen Bjarni Steintún
Danish	Denmark	20 (11)	26.6 (21-37)	Carsten Levisen
Norwegian	Norway	20 (10)	28.4 (20-52)	Åshild Næss
Swedish	Sweden	20 (10)	27.2 (19-61)	Susanne Vejdemo Magnus Enquist
English	England	20 (9)	22.3 (19-31)	Linnaea Stockall Euphemia Snell
Dutch (BE)	Belgium	20 (11)	19.9 (18-26)	Stefanie Fauconnier Jean-Christophe Verstraete, Bram Fauconnier
Dutch (NL)	Netherlands	21 (16)	21.5 (19-27)	Wendy van Ginkel Fiona Jordan Michael Dunn Asifa Majid
Frisian	Netherlands	23 (13)	19.5 (17-24)	Pieter Duijff Arjen Versloot
Luxembourgish	Luxembourg	21 (18)	25.7 (20-40)	Claudine Moulin Claudine Hamen Robert Clees
German	Germany	20 (10)	21.1 (20-25)	Cornelia van Scherpenberg Michael Dunn Fiona Jordan Asifa Majid
Schwyzerdütsch*	Switzerland	20 (10)	25.6 (19-37)	Raphael Berthele Martina Zimmerman

Table 1: The data for the Germanic languages was collected and coded by an expert local team in each country. Without their expertise this project would not have been possible. Data was collected from at least 20 people in each language, of roughly comparable ages.

*Note: Schwyzerdütsch is also known as Swiss German

Design, Stimuli, and Procedure

The project was designed to collect primary speaker data from 4 domains: colour, body parts, containers, and spatial relations. Each participant was tested on all 4 domains; within languages the order of domain was counter-balanced to alleviate fatigue and practice effects. Each session began with a description of the study; information about the procedure and process, and reimbursement (payment was determined by local rates); and written consent was obtained from the participant. At the end of the testing session, all participants completed a full biographical questionnaire which included questions regarding various demographic factors and their language background.

Researchers in each country were given the EoSS Procedure Manual (Majid, Jordan, & Dunn, 2010), which provided detailed instructions on the running conditions, including all the preparations required before testing. All participants were tested in the target language, i.e., Dutch speakers were tested in Dutch; Swedish speakers in Swedish. The full set of instructions was translated and back-translated by the local researchers, and checked by the EoSS office for discrepancies from the Procedure Manual. Any departure from the protocol was resolved by discussion with the local research team. The testing sessions were all audio-recorded for later coding. All materials are archived centrally by the EoSS central office.

Colour

The colour elicitation task consisted of 84 Munsell colour chips arranged in a single fixed random order (materials developed by Majid & Levinson, 2007). The stimuli were in a grey box and presented individually on individual numbered grey plates. Four chips were achromatic. The remaining 80 varied in hue, brightness and saturation, with 20 hues at 4 degrees of brightness. Following the colour naming task, participants also provided data for the best exemplars of each “basic” colour terms in the language (determined by prior elicitation). The Focal Colour task had

the same 84 Munsell colours from the colour naming task but organised in a two-dimensional array according to hue and lightness. The four achromatic colours were presented on the left-hand side. Each colour was uniquely identified by a letter (A-D) and a number (1–20) (Majid, 2008). Participants were asked to indicate the best example of each colour by pointing to it, and the letter-number identifier was recorded. All participants were screened for colour-blindness using 9 colour plates (Waggoner, 2002). Participants were not screened beforehand, but colour data from colour-blind participants are not used in the analyses presented here.

Body parts

Body part terms were elicited using line drawings of a human body. Participants were presented with 90 drawings; 70 pictures were of a human body viewed from the front or back; 20 pictures depicted the head/face from the front. The head pictures followed the body pictures in a single block. Each drawing had a red dot on a location on the body, and the participant was asked to name the body part marked by the dot. The stimulus set was developed specifically for this project (Jordan, Dunn, & Majid, 2009).

Containers

In order to elicit container names, we used the dishes stimuli set developed by Ameel et al. (2005), thanks to the kind permission of the authors. This consisted of 67 photographs of common household dishes, such as bowls, cups, and mugs. The objects were all photographed against a neutral background, and from a constant distance to preserve relative size information. A ruler was also included in the forefront of the photograph, as a further cue to relative size.

Spatial relations

Finally, spatial relation terms were elicited using the Topological Relations Picture Series originally developed by Bowerman and Pederson (1992). In our version, participants were presented with a booklet of 71 line drawings depicting a Figure (in orange) with respect to a

Ground (in black). Participants were asked *Where is the X?* (in their native language), where X was the name of the Figure in each picture.

Coding

The local research team transcribed all participant responses into a standardised coding sheet. Only responses directly relevant to the stimuli were transcribed: conversational asides, commentaries, disfluencies, clear mistakes, etc. were not noted. For each domain we were interested in the “main” responses, which meant that the “full” responses had to be further processed.² Below we summarise the main coding guidelines.

Colour

For the colour task, the main response was the colour term discarding hedges or modifiers. So, *light blue*, for example, was coded as solely *blue*. Source-based terms, such as *mint*, were coded when they were the only response. If however, a source-based term was used to modify a basic colour term, it was not coded: e.g., *mint green* was coded simply as *green*. In cases where more than one basic colour term was used, e.g. *blue-green*, both were coded.

Body parts

For body parts, we coded all conventionalised body part terms. So, for example, in English *forearm* is a complex but conventionalised body part term, whereas *upper shoulder* is not. Further specification of location using descriptors such *upper*, *lower*, *middle*, *behind*, *front* etc. were generally discarded. Similarly, *left* and *right* were not coded as part of the main response. If the body part being indicated was a complex expression that could not be decomposed to indicate the self-same part, e.g. *between the eyes* (*eye* does not indicate the same

² The data presented here are a snapshot from 24-06-2014.

body part as *between the eyes*), the whole description was coded as the main response. These sorts of responses were infrequent.

Containers

For the containers task, the main response was the noun used for each stimulus. Modifiers were discarded; for example the full response *wooden bowl* was coded simply as *bowl*. If participants provided multiple responses, they were all coded. We coded as distinct forms terms that appeared with diminutives (e.g., in Dutch *schaal* and *schaaltje*), departing from Ameel et al. (2005) who collapsed this distinction.

Spatial relations

Researchers transcribed the full responses participants gave (e.g. *the cup is on the table*, *the apple is in the bowl*, *the ribbon is around the candle*). From this, we coded as the main response the preposition (e.g. *in*, *on*, *around*). Complex prepositions (e.g. *on top of*) were coded as a single expression. If there was no identifiable preposition, the main response was coded as missing. Misconstruals of pictures were also discarded (e.g., where the participant did not correctly identify the Figure or Ground).³

Results

We first wanted to determine the similarity of each language to the others for each of the domains under study. In order to measure similarity, the data need to be transformed into the same quantitative metric and so we converted the language data into similarity matrices (cf.

³ We are aware that there may be additional information about spatial location given in other parts of speech (e.g. verbs, case-marking) which are extremely interesting in their own right. For now we leave these complexities aside.

Majid, Boster, & Bowerman, 2008). For each participant, we determined whether the same term was used across each pair of stimuli. If they were, then a 0 was coded, and if not a 1. So, for containers, for example, we have a 67 by 67 matrix (the number of container stimuli in the study), for each individual in each language. These individual matrices were summed to produce a single aggregate matrix for each domain in every language, so that we can compare how similar the languages are to one another overall.

Which domains show the most similarity across the Germanic languages?

In order to quantify how similar the languages were to one another, we compared the similarity matrices of the languages in two different ways. First, we conducted simple Pearson correlations on all possible pairwise similarity matrices, following Malt et al. (1999). Table 2 shows, for each domain, the smallest and largest correlation scores between languages, as well as the average correlation across all languages. From this it is clear that across all domains, these 12 languages share a lot of common structure. All languages are significantly positively correlated (at $p=.01$ level, one-tailed). Nevertheless, some domains appear to share more common structure than others: both colour and body parts correlate on average over 0.9, followed next by containers 0.8, and then spatial relations 0.7.

	Colour	Body parts	Containers	Spatial relations
Minimum r	0.89 Danish-English	0.87 Luxembourgish-Swedish	0.71 Dutch(NL)-Icelandic	0.55 Dutch(BE)-Icelandic
Maximum r	0.96 Dutch(NL)-Frisian Dutch(NL)-Dutch(BE)	0.98 Danish-Norwegian Dutch(BE)-Luxembourgish	0.93 German-Luxembourgish	0.96 Dutch(NL)-Frisian
Average r	0.93	0.94	0.82	0.73

Table 2: Minimum, maximum, and average correlations across all 12 Germanic languages for each domain. The pairs of languages that correlate highest and lowest are given directly under the correlation values (under Minimum r and Maximum r). In cases where there was a tie for the highest/lowest pairs, both language pairs are given.

Next, the languages were submitted to a factor analysis using principal components extraction in SPSS. This technique reduces the data to a smaller number of factors or components where we can measure the overall similarity of the languages by the amount of variance explained by each component (Majid et al., 2008). For all four domains, a single factor accounted for a substantial amount of the variance. This confirms the languages are very similar to one

another. By examining the amount of variance accounted for by the first factor for each domain, we can see the relative similarity of semantic structure across domains. Once again we see the same overall picture: the first component accounted for most variance for colour (93.8%) and body parts (94.5%), followed by containers (83.9%), and then spatial relations (75.8%).

The shared structure is also evident in a separate statistic from the analyses – the eigenvalue scores. For all domains, the first eigenvalue was substantially larger than any of the eigenvalues for subsequent components. In fact, for three of the domains only the first component had an eigenvalue greater than 1, suggesting a single component was sufficient to account for the data. The single exception was spatial relations which required two.

Figure 1 is a visual representation of the factor analyses. It shows the overall similarity of the languages to one another in 4 panels, one for each domain. Languages are plotted together on the basis of how similar they are: the closer the distances, the more similar; the further apart, the more different. In all of the plots the languages load high on the right hand side. This is because the languages all correlate positively with each other: they are making similar distinctions.

Both the correlation analyses and the factor analyses reveal the same result. Consistent with the Gentner hypothesis, and contrary to the Talmy perspective, the semantics for containers are more similar across languages than those for spatial relations. However, and perhaps surprisingly, the most similarity between languages was for colour and body parts. Below we discuss each of the domains in more detail, outlining some interesting points of similarity and difference.

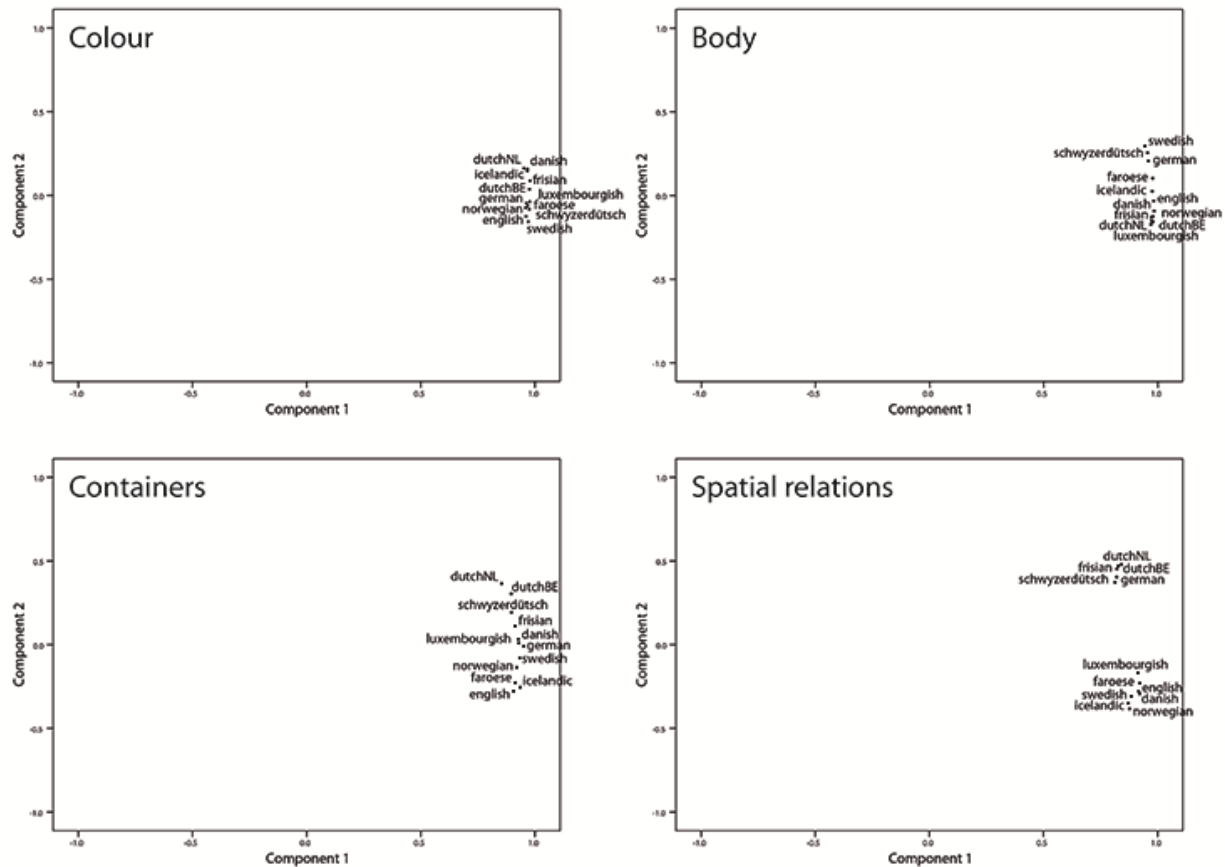


Figure 1: The overall similarity of the Germanic languages to one another for (a) colour, (b) body parts, (c) containers, and (d) spatial relations. The closer two languages are plotted to each other the more similar their lexical semantics for that domain.

Colour

It is clear that there is considerable similarity in the forms used for colour terms in Table 3, as well as the extensional meaning, as illustrated in Figure 2. Each chip in Figure 2 represents a chromatic colour chip used in the colour naming task. The chips that were labelled by the same term are coloured in the same hue; the hue in the figure is an approximation of the best example chosen by participants in the focal colour task (see Methods). There are major similarities, but

small loci of differences too. Note, for example, the subtly differing extensions of orange, yellow, red, and pink, across the languages. On average, there were 33 distinct terms used to describe the colour stimuli, with Faroese speakers using the fewest terms (N=21) and English speakers making the most distinctions (N=54). Table 3 shows the most frequent terms attested in the naming task.

Within the colour domain, English and Danish differed most from one another (see Table 2 with correlations). English *red* is very localised: it was the dominant response for only one colour chip. In contrast, Danish *rød* was a dominant response for 11 different chips. In English, speakers preferentially called most of those chips *pink*, which was then also extended into an area that Danish speakers called *lilla*. Similarly, English *orange* was smaller in extension than its Danish equivalent *orange*. English *yellow* and Danish *gul* are approximately the same size, but because Danish *orange* has a larger extension *gul* seeps into territory that English speakers preferentially name *green*.

Frisian



Dutch (Netherlands)



Dutch (Belgium)



Luxembourgish



German



Schwyzerdütsch



Figure 2: Division of the colour space by modal terms in 12 Germanic languages. Cells coloured the same were referred to by the same term by the majority of speakers (the actual shading of the cell in this figure is a mnemonic only). Where two terms were used equally often, both colours are displayed.

Icelandic	Faroese	Norwegian	Swedish	Danish	English
grænn [546] blár [379] fjólublár [237] bleikur [210] brúnn [173] gulur [145] grár [99] rauður [92] appelsínugulur [75] hvítur [35] lilla [30]	grønt [457] blátt [346] lilla [217] ljósareytt [142] brúnt [95] reytt [88] gult [79] appilsingult [61] grátt [60] violett [40] turkis [38] húðfarvað [35] hvítt [24]	grønn [482] blå [296] lilla [255] rosa [174] brun [140] gul [83] turkis [73] rød [70] oransj [61] grå [53] svart [41] hvit [34] fiolett [21]	grön [427] blå [315] lila [237] rosa [182] brun [117] gul [105] röd [90] turkos [89] orange [77] grå [62] svart [44] vit [39]	grøn [430] blå [279] lilla [221] rød [205] gul [106] brun [100] orange [80] grå [54] turkis [53] hvid [48] sort [47] pink [31] hudfarvet [25]	green [436] blue [290] purple [220] pink [162] brown [105] yellow [91] orange [73] grey [56] turquoise [45] red [33] peach [29] white [28]
Frisian	Dutch (NL)	Dutch (BE)	Luxembourgish	German	Schwyzerdütsch
grien [506] blau [386] pears [266] rôs [222] brún [154] oranje [108] giel [84] griis [71] read [66] swart [35] wyt [31]	groen [465] blauw [349] roze [228] paars [213] bruin [129] geel [105] oranje [85] rood [55] grijs [51] wit [37] lila [20]	groen [453] blauw [318] paars [239] roze [194] bruin [114] geel [101] oranje [78] grijs [70] rood [49] wit [43] zwart [39]	gréng [449] blo [318] mauve [215] rosa [158] brong [126] rout [88] giel [84] orange [74] turquoise [70] schwaarz [59] gro [58] wäiss [25] pink [22]	grün [397] blau [283] lila [209] braun [142] türkis [124] rosa [103] gelb [98] rot [89] orange [84] grau [69] pink [68] violett [38] ocker [36] weiß [30] hautfarben [27] schwarz [24]	grüen [433] blau [265] violett [196] rosa [144] rot [142] bruun [122] türkis [102] gäub [73] oransch [69] grau [53] lila [50] schwarz [38] pink [32] wiiss [25]

Table 3: The most frequent colour names for the 12 Germanic languages, ordered according to their frequency. All terms elicited more than 20 times by the naming task are listed.

Body parts

Like the colour domain, the parts of the body were named in very similar ways across the Germanic languages. Luxembourgish and Swedish were the least similar statistically, but even these languages are strongly correlated with each other (see Table 2). Thus, knowledge of the body-part distinctions made in English would get you pretty far in the other Germanic languages too. Overall, the body task received the most number of distinct terms across all our naming tasks; on average 94 distinct terms were used. Netherlands Dutch had the fewest number of terms (N=62) and Schwyzerdütsch used the most (N=133).

To illustrate the patterns of naming we find, let's examine how speakers of the Germanic languages referred to Figure 3. Speakers of all languages used a form deriving from the Proto-Germanic word **fōt-* 'foot'. In fact, for all the languages, apart from Icelandic, this was the dominant term used for the 4 foot stimuli referred to in Table 4. In Icelandic the term *rist* 'instep' was used twice as often as *fótur* for these stimuli. The cognate term *vríst* in Norwegian, Swedish and Danish also appears with similar extension, but at much lower frequency (English *wrist* is also cognate, but refers to part of the upper limb rather than the lower; see also German *Rist* 'wrist, instep', Old Frisian *handwirst* 'wrist', *fōtwirst* 'instep'; OED).

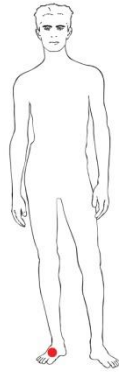


Figure 3: Stimulus number 13 from the EoSS body part stimuli

The semantic extensions of the Germanic ‘foot’ terms were also very similar across the languages, but Frisian looks odd in this comparative perspective. *Foet* in Frisian was extended to name the lower leg and thigh as well. Based on a detailed examination of the Frisian dialects, Versloot (1994) argues that this extension is a result of pejoration of the leg term in Frisian. The leg term comes to be used for animal legs, rather than human ones, and so *foet* becomes a neutral way to refer to the human variety. The Old Norse term *fōtr* also refers to the entire limb (Buck, 1949, p. 242); this is also present in the modern Swedish variety Övdalian (Steensland, 1986, p. 41). There are likewise traces of this in Icelandic: the term *fótur* is used occasionally for lower leg and knee, and there is also the term *fótleggur* ‘leg’, which is a morphologically complex word combining ‘foot’ and a term originally referring to the hollow bone of the limb (there is also *handleggur* ‘arm’).

Icelandic	Faroese	Norwegian	Swedish	Danish	English
rist [51] fótur [25] löpp [3] jarki [3] il [1] tá [1]	fótur [75] fótablað [4] bibbur [1]	fot [67] vrist [9] bein [4] tå [1]	fot [73] fotvalv [3] fotknöl [2] vrist [2] ankel [1]	fod [72] ankel [4] vrist [2] tå [1] hæl [1]	foot [73] ankle [4] toe [2] heel [1]
Frisian	Dutch	Flemish	Luxembourgish	German	Sw. German
foet [75] wreef [9] ankel [4] middenfoets- beentje [3] krop [1]	voet [80] enkel [2]	voet [79] wreef [1] enkel [1]	Fouss [84] Knöchel [1]	Fuß [80] Knöchel [1]	fuèss [54] fuèssrischt [11] fuèssruggè [6] rischt [5] mittufuèss [2] fuèssglènk [1] ungèrhaub vom fuèss-chnöchu [1]

Table 4: Terms used to name the stimuli left front foot, right front foot, left back foot, right back foot.

Unlike ‘foot’, there is no single cognate set used in the Germanic languages for ‘leg’, and we see a more divergent pattern of naming. Participants saw red dots on the upper and lower leg, on the right and left side of the body, viewed from front and back. This lets us examine how they chose to refer to these distinct images. English speakers predominantly used *leg* for all such stimuli; similarly Dutch, Flemish and Luxembourgish used a single term *been/Been*. Frisian speakers also used *been*, but not as frequently. However, the speakers of the other languages tended to use distinct terms for the upper and lower leg. German speakers used *Oberschenkel* and *Unterschenkel*, literally ‘upper-leg’ and ‘lower-leg’, but the root **Schenkel* alone was never attested. Schwyzerdütsch speakers used the cognate terms, *obèrschènku* and *ungèrschènku*, although bare *schènku* was also used (but only twice, and only with reference to upper leg

stimuli). The Scandinavian languages make sporadic use of *bein/ben*, but for the most part other terms were more frequent. Like German speakers, the participants preferentially distinguished upper and lower leg. This discussion illustrates that even in a domain with much shared structure, there are still many intricacies to be observed.

Containers

As the previous analyses showed, there is substantial agreement across the languages in naming patterns for colour and body parts but things begin to differentiate further for containers, where the first factor in the factor analysis only accounted for around 76% of the data. That means that there are points of disagreement across the languages in how various household items ought to be categorised for the purposes of speaking about them. Across the languages, on average 39 different terms were used to describe the container photographs. Faroese used the least number of distinct terms (N=19), whereas Schwyzerdütsch used the most (N=71). Table 5 provides a list of all the terms that were produced at least 20 times in each language, ordered by their rank frequency.

As an illustrative example we focus briefly on the naming of cups, mugs, and glasses as shown in Table 6. For each language, we plot the modal term (the term used most often) for each of the stimuli depicted in the top row. The extension of each term can then be read off the rows for each language. The 12 languages use similar forms to refer to related objects, but the extensions of the terms differ in an ordered way. Faroese speakers used the term *koppur* to refer to small cups, larger mugs, vessels for drinking beer, and glasses without any handles. Dutch speakers in Belgium used a general term *tas* for all the drinking vessels with a handle, but used *beker* for the vessel without a handle. Frisian speakers, on the other hand, used *beker* for almost all of the stimuli in Table 6, reserving *kopke* for the small vessels with handles. Danish makes the

distinction in a slightly different place again with *kop* for cups and mugs, but *krus* for tankard and glass.


The most common naming strategy, exemplified by Norwegian, Icelandic, Luxembourgish, German and Schwyzerdütsch, is to make a three-way distinction: one term for cups and mugs, a second term for tankard, and a third term for glass. In Swedish a three-way distinction cuts the space slightly differently, with *kopp* used for small vessels with a handle and middling sized wide vessels with handles. Taller vessels with handles, including tankard, are called *mugg*, with *glass* reserved for the vessel without a handle. The Swedish forms are similar to those in English, but English speakers make a different distinction again. Speakers used *cup* for smaller vessels with a handle, but 17 speakers also used *cup* for the glass stimulus (two speakers called it a *beaker*, and only one speaker called it a *glass*). *Mug* was the modal response for the tankard, and it tied with *cup* for the wide vessel with a handle. Netherlands Dutch speakers made the most distinctions, distinguishing smaller cups (*kopje*), from the wider vessel (*kom*), taller vessels with handle (*mog*), and the glass (*beker*).

Malt and colleagues have illustrated that languages can vary substantially in their lexical categories for containers (Ameel, Malt, & Storms, 2008; Ameel et al., 2005; Malt et al., 1999; Pavlenko & Malt, 2013). We have shown here that even very closely related languages are making subtly different distinctions. The two varieties of Dutch in our sample illustrate this neatly. Belgian Dutch speakers used only two terms, but Netherlands Dutch speakers made a four-way distinction between the drinking vessels in Table 6. In fact, the general term *tas* used for most of the drinking vessels by Belgian Dutch speakers has a completely different meaning in the Netherlands, namely, ‘bag’ or ‘satchel’. So don’t make the mistake of asking for a *tas* of coffee when visiting Amsterdam!

This variation has to be appreciated against the backdrop of similarity that we saw in the previous section. Figure 1 shows that overall in this domain, Netherlands and Belgian Dutch are highly similar. Although the drinking vessels are an illustrative locus of variation, many of the stimuli depicted dishes and bowls, where the similarities appear to be larger than is apparent in Table 6.

Icelandic	Faroese	Norwegian	Swedish	Danish	English
skál [509] bolli [287] diskur [166] fat [97] glas [69] kanna [59] mót [56] krús [31] öskubakki [30] mál [29] ílát [25] undirskál [25] bakki [24]	skál [503] koppur [273] talerkur [123] krúss [110] fat [105] ílát [103] glas [53] bolli [24]	skål [295] kopp [292] bolle [250] fat [126] tallerken [101] krus [88] form [79] glass [71] asjett [31]	skål [428] kopp [210] fat [181] mugg [149] tallrik [147] form [84] glas [62]	skål [413] kop [248] fad [200] tallerken [189] krus [141] glas [49] askebæger [36]	bowl [462] cup [203] mug [174] dish [157] plate [128] glass [53] ashtray [40] pot [28] tray [20]
Frisian	Dutch [NL]	Dutch [BE]	Luxembourgish	German	Schwyzerdütsch
skaaltsje [300] kopke [191] skaal [178] pantsje [118] beker [113] kom [110] bakje [104] panne [78] glês [75] board [70] mok [63] kop [26]	schaal [277] schaaltje [194] bord [132] kopje [122] bakje [119] mok [111] kom [68] beker [66] glas [59] kommetje [56] bordje [35] kop [30] pul [28]	kom [321] bord [235] schaal [190] tas [185] glas [79] beker [58] kop [54] schotel [46] pot [38] mok [32] potje [24]	Schossel [500] Taass [291] Teller [194] Becher [55] Glas [44] Dëppchen [37] Plateau [36] Béierkrou [29] Form [23] Äschebecher [21] Schuel [21] déiwen Teller [21] Humpen [20]	Schale [344] Tasse [266] Schüssel [199] Teller [197] Becher [76] Form [57] Krug [55] Glas [54] Platte [23]	schalè [202] täuèr [196] tassli [172] schäli [160] tassè [111] schüssèli [63] glas [59] form [55] schüsslè [35] bächèr [31] humpè [31] täuèrli [30] plattè [25]

Table 5: The most frequent container names for the 12 Germanic languages, ordered according to their frequency. All terms elicited more than 20 times by the naming task are listed.



<i>Faroese</i>	koppur				
<i>Dutch (BE)</i>	tas				beker
<i>Frisian</i>	kopke		beker		
<i>Danish</i>	kop			krus	
<i>Norwegian</i>	kopp			krus	glass
<i>Icelandic</i>	bolli			krús	glas
<i>Luxembourgish</i>	Taass			Béierkrou	Becher
<i>German</i>	Tasse			Krug	Becher
<i>Schwyzerdütsch</i>	tassli			humpè	bächèr
<i>Swedish</i>	kopp		mugg		glass
<i>English</i>	cup		mug		cup
<i>Dutch (NL)</i>	kopje		kom	mok	beker

Table 6: Naming patterns for drinking vessels across the Germanic languages. (Capitalization of the nouns in German and Luxembourgish follows the conventions of those languages' writing systems.)

Spatial relations

The semantics of spatial prepositions were the most variable across the Germanic languages. The spatial relations task received on average 33 distinct terms (comparable to the number of terms in the colour naming task). Netherlands Dutch had the fewest number of terms (N=20) and Faroese used the most (N=54). In Table 2, we saw that Dutch and Icelandic correlated least with each other. In Figure 1 Dutch and Icelandic are at opposite ends of the y-axis, reiterating the findings from the correlational analyses. This figure, in fact, illustrates there are two clusters of languages for spatial relations. English, Faroese, Danish, Norwegian, Swedish, Luxembourgish and Icelandic sit in one cluster, whilst Netherlands Dutch, Belgian Dutch, Frisian, Schwyzerdütsch, and German sit together in another cluster. The major distinctions between the languages can be illustrated (in a simplified way) in Table 7. The table is adapted

from Bowerman and Choi (2001), who singled out these scenes⁴ to illustrate differences between diverse languages. In our sample of Germanic languages, it seems we have two distinct semantic strategies: there are the English-type *on* languages and Dutch-type *op-aan* languages.

Faroese, Danish, Norwegian, Swedish, and Icelandic⁵ are like English in that they have one preposition that can be used for cup-on-table, bandaid-on-skin, apple-on-twigg, and handle-on-cupboard. Frisian, Schwyzerdütsch, and both varieties of Dutch, on the other hand, make a distinction between these scenes. Bowerman (1996) suggested that the distinction between *op* and *aan* has to do with the force-dynamics of the situation. It is an *aan*-type relation if the speaker conceptualised the Figure as pulled-on by a force (e.g., gravity) that has to be counteracted in order to maintain contact with the Ground. It is an *op*-type relation if the Figure is taken to be resting comfortably on the Ground (with no counteracting forces). In earlier work, Bowerman suggested that the *op-aan* distinction was rare and that it only occurred in Dutch (Bowerman & Choi, 2001, p. 484), but we see here that Frisian and Schwyzerdütsch exemplify it too. As also noted by Bowerman, German makes a somewhat similar distinction but, as she noted, it differs in some essential respects. *Auf* in German is more restricted than Dutch *op*, and is not used for situations such as bandaid-on-skin; nor is it used for insect-on-ceiling, rain-on-windowpane, butter-on-knife, etc. *Auf* seems to have a stronger requirement for support from below.

⁴ Plus one additional scene we do not consider here for simplicity.

⁵ In Icelandic *ofan á* was actually the dominant response for cup-on-table. But for every stimulus that *ofan á* can be used *á* is also used, and it is *also* usually the dominant response. In fact, *ofan á* is only the dominant response for cup-on-table (13 speakers used *ofan á* vs 7 speakers used *á*). Therefore, for simplicity, we use the widest applicable form in the table.

Luxembourgish also makes a contrast between different “on” relations but the cut is somewhat different again. It only makes an explicit contrast between different “on” relations where the Ground object is grammatically feminine, but collapses the contrast when the Ground is masculine. Where the Ground object is a definite noun with feminine gender then speakers make a distinction between *op* and *un*. However when the Ground object is expressed as a noun with masculine gender then speakers use the spatial preposition *um* in a manner analogous to English *on*. This is a quirk of the grammar, however, it does beg the question of how speakers are representing these relations non-linguistically. We leave this as a matter for future investigation.

	cup on table	bandaid on skin	apple on twig	handle on cupboard	apple in bowl
English	on				in
Faroese	á				í
Danish	på				i
Norwegian	på				i
Swedish	på				i
Icelandic	á				í
Luxembourgish	op/um		un/um		an
Dutch (NL)	op		aan		in
Dutch (BE)	op		aan		in
Frisian	op		oan		yn
Schwyzerdütsch	uf		a		i
German	auf	an			in

Table 7: Prepositional use across the Germanic languages for 5 illustrative scenes from the spatial relations stimuli.

Conclusions

In the opening section we asked whether some semantic domains are more similar across languages than others. According to the Gentner prediction, containers ought to show the most similarity and spatial relations show the least. The other two domains fall out of the framework

proposed by Gentner, but based on the underlying rationale one could predict colour and body parts should fall somewhere between containers and spatial relations. In contrast, the perspective favoured by many linguists (notably Talmy) makes the opposite prediction: spatial relations should display more regularity cross-linguistically than containers, colours, or body parts. Our analyses of the Germanic languages do not fit the Talmyan perspective, but rather are more compatible with the Gentnerian perspective. There was more similarity between languages in the semantics of containers than spatial relations. However, colours and body parts showed most similarity overall. This was unexpected in relation to current theorising about semantics, and raises interesting questions for future research, such as: is such high degree of semantic similarity peculiar to just these two domains, or is this a more general property of the semantics of attributes and parts? Would we see the same similarities in meanings across languages if we were to test tastes, smells, and patterns, for example?

Going back to the four different domains under study within EoSS – colours, body parts, containers, and spatial relations – we found substantial similarities in meanings across the Germanic languages. From the current analyses we cannot say whether these shared meanings are the result of universal cognitive principles, common linguistic~cultural heritage, or on-going contact and exchange between these communities. But this is something that we can answer in the future using the approaches we outlined in the introduction. This promises additional new insights into the patterns and processes of historical change for different semantic categories.

Against this backdrop of similarity, we also saw subtle semantic variations played out in different ways in each speech community. Danish *rød* encompasses a larger swathe of colour space than English *red*. Frisian *foet* extends to the lower leg and thigh, unlike its cognate in the other Germanic languages. Belgian Dutch speakers use *tas* for drinking vessels with a handle, but in the Netherlands variety of Dutch *tas* means ‘bag’, and speakers must choose between *kopje*,

kom, or *mok* to refer to the same drinking vessels. And support relations are treated differently across the Germanic languages, with some languages making a distinction between situations such as cup-on-table and handle-on-cupboard (e.g., Dutch), whereas other languages treat these situations as equivalent (e.g., English). In each language, in each domain the language learner must become acquainted with the local conventions of semantic packaging to become a competent speaker~hearer.

The challenges of learning a specific semantic system, even in closely related languages, is exemplified in the other papers of this volume too. Vejdemo and colleagues (2014) examine how ‘pink’ is lexicalised in different Germanic languages, while Zimmermann and colleagues (2014) explore how ineffable hues are handled in some Germanic languages by the term ‘skin-coloured’. Levisen (2014) zooms in on the lexicalisation of ‘neck’ and ‘throat’ in Scandinavian languages; while Whelpton and colleagues (2014) compare container naming strategies in Icelandic and Dutch; and Berthele and colleagues (2014) discuss the intricacies of spatial semantics in the Germanic languages. Each of these papers provide detailed information on specific forms and their functions, and thus illustrate the complexities disguised by the generalisations provided in our introduction.

The papers also illustrate different methodologies for contextualising the snapshot extensional approach we utilise within the EoSS project. For example, Whelpton et al. and Berthele et al. explore the relationship between morphosyntax and semantics. Both underline the importance of understanding the relationship between lexical packaging and combinatorial possibilities. This is further illustrated by Levisen who uses phraseology, or conventionalised phrasemes, such as “bottom of my heart” as critical data in his semantic analyses. Levisen takes the audacious step of using his extensional data from the EoSS task as the basis for a Natural Semantic Metalanguage (e.g., Wierzbicka, 1996) intensional analysis. The results are thought-

provoking. Vejdemo et al. use a different approach again. They incorporate a historical dimension in their analyses appealing to both linguistic and cultural factors as means of understanding current language use. Zimmermann et al. appeal to a different framework again, using sociolinguistics and insights from race theory to understand colour semantics. These myriad approaches prove that extensional data can be incorporated into disparate frameworks. We hope they provide inspiration for other studies of the Indo-European languages in the future.

Acknowledgements

This work stems from the Evolution of Semantic Systems project and received financial support from the Max Planck Gesellschaft. We would like to thank our Germanic collaborators who are listed by name in Table 1. Without their expertise this project would not have been possible. We would also like to thank Carolin Lorenz, Wendy van Ginkel, Tuong-Van Vu for their tireless efforts in data preparation and data checking. Thanks to Ludy Cilissen for assistance with illustrations. Finally, we would like to thank Nigel Love, Umberto Ansaldi, Dan Dediu and Barbara Malt for critical feedback on earlier versions of this paper.

References

- al-Jehani, N. M. (1990). Color terms in Mecca: A sociolinguistic perspective. *Anthropological Linguistics*, 32(1/2), 163–174.
- Ameel, E., Malt, B., & Storms, G. (2008). Object naming and later lexical development: From baby bottle to beer bottle. *Journal of Memory and Language*, 58(2), 262–285.
- Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language*, 53(1), 60–80. doi:10.1016/j.jml.2005.02.004

- Andersen, E. S. (1978). Lexical universals of body-part terminology. In J. H. Greenberg, C. A. Ferguson, & E. A. Moravcsik (Eds.), *Universals of human language* (Vol. 3, pp. 335–368). Stanford, CA: Stanford University Press.
- Bazell, C. E. (1953). *Linguistic form*. Istanbul: Istanbul Press.
- Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley: University of California Press.
- Berthele, R., Whelpton, M., Næss, Å., & Duijff, P. (2014). Static spatial descriptions in five Germanic languages. *Language Sciences*. doi:10.1016/j.langsci.2014.07.006
- Bloomfield, L. (1933). *Language*. New York, NY: Henry Holt and Company.
- Bouckaert, R., Lemey, P., Dunn, M., Greenhill, S. J., Alekseyenko, A. V., Drummond, A. J., ... Atkinson, Q. D. (2012). Mapping the Origins and Expansion of the Indo-European Language Family. *Science*, 337(6097), 957–960. doi:10.1126/science.1219669
- Bowerman, M. (1992). Topological relations pictures: static relations. In S. C. Levinson (Ed.), *Space stimuli kit 1.2: November 1992* (pp. 25–28). Nijmegen: Max Planck Institute for Psycholinguistics.
- Bowerman, M. (1996). The origins of children's spatial semantic categories: Cognitive versus linguistic determinants. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 145–176). Cambridge: Cambridge University Press.
- Bowerman, M., & Choi, S. (2001). Shaping meanings for language: Universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman & S. C. Levinson (Eds.), *Language Acquisition and Conceptual Development* (pp. 475–511). Cambridge: Cambridge Univ. Press.
- Bowerman, M., & Pederson, E. (1992). Topological relations picture series. In S. C. Levinson (Ed.), *Space stimuli kit 1.2: November 1992* (p. 51). Nijmegen: Max Planck Institute for Psycholinguistics.

- Bowerman, M., & Pederson, E. (1993). Cross-linguistic studies of spatial semantic organization. In P. Brown, G. Senft, & L. Wheeldon (Eds.), (Vol. Max Planck Institute for Psycholinguistics: Annual report 1992, pp. 53–56). Nijmegen: MPI for Psycholinguistics.
- Brown, C. H. (1976). General principles of human anatomical partonomy and speculations on the growth of partonomic nomenclature. *American Ethnologist*, 3(3), 400–424.
- Buck, C. D. (1949). *A dictionary of selected synonyms in the principal Indo-European languages: A contribution to the history of ideas*. Chicago: University of Chicago Press.
- Chomsky, N. (1993). *Language and thought*. Wakefield, RI: Moyer Bell.
- Comrie, B., & Corbett, G. G. (Eds.). (1993). *The Slavonic languages*. London: Routledge.
- Dahl, Ö. (2008). An exercise in a posteriori language sampling. *STUF-Language Typology and Universals Sprachtypologie Und Universalienforschung*, 61(3), 208–220.
- Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories in a stone-age tribe. *Nature*, 398(6724), 203–204.
- De Groot, A. M. B. (1993). Word-type effects in bilingual processing tasks: Support for a mixed-representational system. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 27–51). Amsterdam: John Benjamins.
- De Vignemont, F., Majid, A., Jola, C., & Haggard, P. (2009). Segmenting the body into parts: Evidence from biases in tactile perception. *The Quarterly Journal of Experimental Psychology*, 62(3), 500–512.
- De Vignemont, F., Tsakiris, M., & Haggard, P. (2005). Body mereology. In G. Knoblich, I. M. Thornton, M. Grosjean, & M. Shiffrar (Eds.), *Human Body perception from the inside out* (pp. 147–170). New York: Oxford University Press.
- De Vos, C. (2011). Kata Kolok color terms and the emergence of lexical signs in rural signing communities. *The Senses and Society*, 6(1), 68–76. doi:10.2752/174589311X12893982233795
- Dryer, M. S. (1989). Large linguistic areas and language sampling. *Studies in Language*, 13(2), 257–292.

- Dunn, M., Greenhill, S. J., Levinson, S. C., & Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, 473, 79–82.
- Felsenstein, J. (2004). *Inferring phylogenies*. Sunderland: Sinauer Associates.
- Fortunato, L., & Jordan, F. M. (2010). Your place or mine? A phylogenetic comparative analysis of marital residence in Indo-European and Austronesian societies. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 365(1559), 3913–3922.
- Gentner, D. (1981). Some interesting differences between verbs and nouns. *Cognition and Brain Theory*, 4(2), 161–178.
- Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge: Cambridge University Press.
- Haspelmath, M. (2003). The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (Vol. 2). Mahwah, NJ: Lawrence Erlbaum Associates.
- Haspelmath, M. (2009). Lexical borrowing: Concepts and issues. In M. Haspelmath & U. Tadmor (Eds.), *Loanwords in the world's languages* (pp. 35–54). Berlin: de Gruyter.
- Hill, C. (2011). Named and unnamed spaces: Color, kin, and the environment in Umpila. *The Senses and Society*, 6(1), 57–67. doi:10.2752/174589311X12893982233759
- Jones, S. W. (1798). The third anniversary discourse, delivered 2d February, 1786: on the Hindus. *Asiatick Researches*, 1, 415–431.
- Jordan, F. M., Dunn, M., & Majid, A. (2009). Body Part naming booklet. Developed for the EoSS project. Nijmegen: Max Planck Institute for Psycholinguistics.
- Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: Evidence for “fast mapping.” *Science*, 304(5677), 1682–1683.
- Kay, P., Berlin, B., Maffi, L., Merrifield, W. R., & Cook, R. (2009). *The world color survey*. Stanford, CA: CSLI Publications.

- Kay, P., & Regier, T. (2003). Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences of the United States of America*, 100(15), 9085–9089.
- Landau, B., & Jackendoff, R. (1993). “What” and “where” in spatial language and spatial cognition? *Behavioral and Brain Sciences*, 16(02), 255–265.
- Lazar-Meyn, H. A. (1991). The colour systems of the Modern Celtic languages: Effects of language contact. In *Language contact in the British Isles: Proceedings of the Eighth International Symposium on Language Contact in Europe, Douglas, Isle of Mann, 1988* (pp. 227–241). Tübingen, Germany: Max Niemeyer Verlag.
- Levinson, S. C. (2000). Yélf Dnye and the theory of basic color terms. *Journal of Linguistic Anthropology*, 10(1), 3–55.
- Levinson, S. C., Meira, S., & The Language and Cognition Group. (2003). “Natural concepts” in the spatial topological domain-adpositional meanings in crosslinguistic perspective: an exercise in semantic typology. *Language*, 79(3), 485–516.
- Levinson, S. C., & Wilkins, D. (Eds.). (2006). *Grammars of Space: Explorations in Cognitive Diversity*. Cambridge: Cambridge University Press.
- Levisen, C. (2014). Scandinavian semantics and the human body: an ethnolinguistic study in diversity and change. *Language Sciences*. doi:10.1016/j.langsci.2014.05.004
- Lewis, P. M., Simons, G. F., & Fenning, C. D. (Eds.). (2014). *Ethnologue: Languages of the world* (Seventeenth edition.). Dallas, Texas: SIL International. Retrieved from <http://www.ethnologue.com>
- Lindsey, D. T., & Brown, A. M. (2004). Sunlight and “blue”: The prevalence of poor lexical color discrimination within the “grue” range. *Psychological Science*, 15(4), 291–294.
- Lindsey, D. T., Brown, A. M., Brainard, D., & Apicella, C. (2014). Hadza color naming and the origins of basic color categories. *Journal of Vision*, 14(10).
- Lucy, J. A. (1997). Linguistic relativity. *Annual Review of Anthropology*, 26(1), 291–312. doi:10.1146/annurev.anthro.26.1.291

- Mace, R., Holden, C. J., & Shennan, S. J. (Eds.). (2005). *The evolution of cultural diversity: A phylogenetic approach*. London: UCL Press.
- Mace, R., & Pagel, M. (1994). The comparative method in anthropology. *Current Anthropology*, 549–564.
- Majid, A. (2008). Focal colors. In *Field Manual* (Vol. 11, pp. 8–10). Nijmegen: Max Planck Institute for Psycholinguistics.
- Majid, A. (2010). Words for parts of the body. In B. C. Malt & P. Wolff (Eds.), *Words and the Mind: How words capture human experience* (pp. 58–71). New York, NY: Oxford University Press.
- Majid, A. (2014). Comparing lexicons cross-linguistically. In J. Taylor (Ed.), *The Oxford Handbook of the Word*. Oxford: Oxford University Press.
- Majid, A., Boster, J. S., & Bowerman, M. (2008). The cross-linguistic categorization of everyday events: A study of cutting and breaking. *Cognition*, 109(2), 235–250.
- Majid, A., Bowerman, M., van Staden, M., & Boster, J. S. (2007). The semantic categories of cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics*, 18(2), 133–152.
- Majid, A., Enfield, N. J., & van Staden, M. (Eds.). (2006). Parts of the body: Cross-linguistic categorisation [Special Issue]. In *Language Sciences* (Vol. 28(2–3)).
- Majid, A., Jordan, F. M., & Dunn, M. (2010). *Evolution of Semantic Systems: Procedures Manual*. Nijmegen: Max Planck Institute for Psycholinguistics.
- Majid, A., & Levinson, S. C. (2007). The language of vision I: colour. In A. Majid (Ed.), *Field Manual Volume 10* (pp. 22–25). Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from <http://fieldmanuals.mpi.nl/volumes/2007/language-of-vision-colour/>
- Majid, A., & van Staden, M. (submitted). Can nomenclature for the body be explained by embodiment theories?, *Topics in Cognitive Science*.
- Malt, B. C., & Majid, A. (2013). How thought is mapped into words. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(6), 583–597.
- Malt, B. C., Sloman, S. A., & Gennari, S. P. (2003). Universality and language specificity in object naming. *Journal of Memory and Language*, 49(1), 20–42.

- Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language*, 40(2), 230–262. doi:10.1006/jmla.1998.2593
- Nichols, J. (1992). *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
- Nunn, C. L. (2011). *The comparative approach in evolutionary anthropology and biology*. Chicago: University of Chicago Press.
- Pavlenko, A., & Malt, B. C. (2013). Kitchen Russian: Cross-linguistic differences and first-language object naming by Russian–English bilinguals. *Bilingualism: Language and Cognition*, 14(01), 19–45. doi:10.1017/S136672891000026X
- Pepperberg, I. M. (2012). Symbolic communication in the Grey Parrot. In J. Vonk & T. K. Shackelford (Eds.), *The Oxford handbook of comparative evolutionary psychology* (pp. 297–319). New York: Oxford University Press.
- Perkins, D. R. (1989). Statistical techniques for determining language sample size. *Studies in Language*, 13(2), 293–315.
- Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences*, 104(4), 1436–1441.
- Regier, T., Khetarpal, N., & Majid, A. (2013). Inferring semantic maps. *Linguistic Typology*, 17, 89–105.
- Rijkhoff, J., Bakker, D., Hengeveld, K., & Kahrel, P. (1993). A method of language sampling. *Studies in Language*, 17(1), 169–203.
- R.L.G. (2013, May 29). Lexical facts. Retrieved from <http://www.economist.com/blogs/johnson/2013/05/vocabulary-size>
- Roberson, D., Pak, H., & Hanley, J. R. (2008). Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition*, 107(2), 752–762.
- Steels, L., & Belpaeme, T. (2005). Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences*, 28(04), 469–489.

- Steensland, L. (1986). *Material till en älvdalsk ordbok : svenska-älvdalska, älvdalska-svenska*. Älvdalen: Ulum dalska.
- Talmy, L. (1983). How language structures space. In H. L. Pick Jr & L. P. Acredolo (Eds.), *Spatial orientation: Theory, research and application*. New York: Plenum Press.
- Turton, D. (1980). There's no such beast: Cattle and color naming among the Mursi. *Man*, 15(2), 320–338.
- Vejdemo, S., Levisen, C., van Scherpenberg, C., Guðmundsdóttir Beck, Þórhalla, Næss, Å., Zimmermann, M., ... Whelpton, M. (2014). Two kinds of pink: development and difference in Germanic colour semantics. *Language Sciences*. doi:10.1016/j.langsci.2014.07.007
- Versloot, A. P. (1994). *“Fuß” und “Bein” in den modernen friesischen Mundarten, eine strukturell-semantische Untersuchung* (MA). Christian-Albrechts-Universität, Kiel.
- Waggoner, T. L. (2002). Quick six colour vision test pseudoisochromatic plates. In *Colour vision testing made easy*. Good-Lite company.
- Weinreich, U. (1980). *On semantics*. (W. Labov & B. S. Weinreich, Eds.). Philadelphia: University of Pennsylvania Press.
- Whelpton, M., Guðmundsdóttir Beck, Þórhalla, & Jordan, F. M. (2014). The semantics and morphology of household container names in Icelandic and Dutch. *Language Sciences*. doi:10.1016/j.langsci.2014.07.014
- Wierzbicka, A. (1996). *Semantics: primes and universals*. Oxford University Press.
- Wierzbicka, A. (2005). There are no “color universals” but there are universals of visual semantics. *Anthropological Linguistics*, 47(2), 217–244.
- Yule, H., & Burnell, A. C. (1903). *A glossary of colloquial Anglo-Indian words and phrases, and of kindred terms, etymological, historical, geographical and discursive*. London: John Murray.
- Zimmermann, M., Levisen, C., Guðmundsdóttir Beck, Þórhalla, & van Scherpenberg, C. (2014). Please pass me the skin coloured crayon! Semantics, socialisation, and folk models of race in contemporary Europe. *Language Sciences*. doi:10.1016/j.langsci.2014.07.011